

Time-Series Classification of High-Temporal Resolution AVHRR NDVI Imagery of Mexico

Stephen L. Egbert, Miguel Ortega-Huerta, Enrique Martinez-Meyer,
Kevin P. Price, and A. Townsend Peterson*

Kansas Applied Remote Sensing Program and Department of Geography

*Museum of Natural History and Department of Ecology and Evolutionary Biology

University of Kansas

Lawrence, KS 66045

Telephone: 785.864.7719 Fax: 785.864.0392

Email: s-egbert@ukans.edu

ABSTRACT

Time-series data from wide-field sensors, acquired for the period of a growing season or longer, capitalize on phenological changes in vegetation and make it possible to identify vegetated land cover types in greater detail. Our objective was to examine the utility of time-series data to rapidly update maps of vegetation condition and land cover change in Mexico as an input to biodiversity modeling. We downloaded AVHRR NDVI 10-day composites from the USGS EROS Data Center for 1992-1993 and adjusted for cloud contamination by further aggregating the data. In the first phase of our analysis, we selected training sites for various land cover types using a land cover map created by the Mexican National Institute of Statistics, Geography, and Informatics (INEGI) as a guide. Since there is a high degree of spectral variability within many of the vegetated land cover types, we subjected the spectral response patterns to cluster analysis. We then used the statistics of the clusters as training data in a supervised classification. We also compared unsupervised and univariate decision tree approaches, but these provided unsatisfactory results. Best results were achieved with a 19-class map of land use/land cover employing a supervised approach.

INTRODUCTION

Creating maps of Mexico's vegetation types and land use patterns quickly and inexpensively is a key component in a larger NSF-funded project focused on modeling the potential consequences of global climate changes on elements of biodiversity in Mexico. Although Landsat MSS, Landsat TM and SPOT have been successfully used to characterize regional land cover types in Mexico in the past, the large area and environmental heterogeneity of the country make mapping its general vegetation types using such high resolution data in an ongoing monitoring program impractical.

Although the spatial resolution of the NOAA AVHRR sensor is coarse (1.1 km), it is in part compensated for by its frequency of repeat coverage, making AVHRR data and derived products highly useful for monitoring land cover change. Specifically, the Normalized Difference Vegetation Index (NDVI) derived from the AVHRR data set is widely

used in monitoring continental and global vegetation distribution and dynamics, drought severity and location, and environmental deterioration [1].

Previous attempts to use AVHRR data to classify Mexico's vegetation types include the work of Sorani and Alvarez [2], who produced a hybrid cartographic product using AVHRR, TM, and ancillary data. Mora and Iverson [3] applied a principal components technique on Global Vegetation Index (GVI) data to generate an unsupervised classification for the country. The latter study took advantage of the phenological characteristics of vegetation revealed by the vegetation index for evaluating spectral responses.

The major goal of this project was to develop a reliable and acceptably accurate classification methodology for mapping Mexico's main vegetation formations at a country-wide scale, making possible the development of assessment and monitoring programs for revealing conditions and changes in land-cover types for the entire country. The present project used time-series AVHRR remotely sensed data, with broad coverage and high temporal resolution. We tested the applicability of AVHRR NDVI data for classifying Mexico's main vegetation types to satisfy the methodological and conceptual needs for modeling and monitoring spatial and temporal patterns of biodiversity.

METHODS

The data used in this project consisted of 10-day NDVI composites from 1992 and 1993, provided by the Global Land Cover 1 km AVHRR Project of the EROS Data Center of the U.S. Geological Survey. To create an NDVI time series and to increase the probability of obtaining images free of cloud contamination, the three 10-day composites for each month were aggregated by selecting the maximum NDVI value for each pixel. The monthly NDVI images were examined visually to select those with minimal apparent cloud contamination. Nine NDVI monthly images were selected as being most cloud free: January 1992, April 1992, June 1992, October 1992, November 1992, February 1993, March 1993, July 1993, and August 1993. These images were layer stacked to create a 9-band image that was used to perform the vegetation classification. We examined supervised, unsupervised, and univariate decision tree classification approaches. The latter two approaches

provided markedly inferior results and were not considered further.

An updated, high resolution land cover/land use map, generated by INEGI using TM data, was used as the reference data for our classification, both for locating training sites and for "ground truth" data. Our vegetation classification followed that proposed by Rzedowski [4]. In addition, four land use classes (cropland, irrigated cropland, human-introduced grassland, and cultivated grassland) were added, for a total of 26 land use/land cover (LULC) classes. INEGI's LULC map, which originally had more than 300 classes, was adapted to our classification, and training areas were digitized for each of the 26 LULC classes across the country. This was done by overlaying INEGI's vector coverage of land cover on the 9-layer NDVI image.

Because of the large number of training signatures gathered and the potential spectral variability of the training sites within each class, we used a two-step approach to refine the signatures to be used in training the supervised classifier. First, for each class we subjected the signatures to a hierarchical cluster analysis. Then we examined each of the resulting clusters by viewing the spectral profiles of the signatures that comprised the cluster in question. Based on the spectral profiles, we eliminated some of the clusters and merged others, resulting in 135 total clusters representing the 26 LULC classes. Some LULC classes were represented by nine or ten clusters (which we termed "subclasses"), while some were represented by only one or two.

The signatures for the 135 subclasses were then used to train a maximum likelihood classifier for the nine-month NDVI time-series image. The output was a 26-class LULC map for Mexico. To assess the classification's performance, we performed a correspondence analysis and created a standard contingency matrix. For the correspondence analysis, we generated 10,000 random points from the classified image and used the INEGI LULC map as reference data to cross-tabulate the results.

Evaluation of the initial classification performance was followed by the inclusion or removal of subclass signatures, which meant an iterative process for exploring and improving the classification. The results of the correspondence analysis provided the basis for deleting, merging, and creating new signatures. Of the 135 signatures generated at the beginning of this process, 104 signatures representing 19 classes were retained to perform the final classification. The final classification was also subjected to a generalization process, with a minimal area size of 10 pixels connected under the 8-pixel neighbor rule. We then repeated the correspondence analysis for the final 19-class map.

RESULTS

One outcome of the iterative process for identifying clusters of similar signatures (i.e., subclasses) was a regionalization of signatures within a single class, demonstrating how the time-series NDVI data captured

regional differences in phenological responses within vegetation formations across Mexico. An example of this is the regional variation in signatures within tropical deciduous forest. One cluster of signatures represents the northeastern region (Tamaulipas and San Luis Potosi), another is located in the southern region towards the Pacific Coast (Oaxaca, Guerrero, and Michoacan), a third runs parallel to the Pacific coast in the northwest (Sonora, Sinaloa, and Nayarit), and one is in the interior south in the Oaxaca highlands.

For the initial 26-class map, an overall correspondence value (CV) of 46.41% ($\kappa = 0.4052$) between the INEGI map and the AVHRR map was obtained. However, the CVs for particular LULC classes varied widely. CVs varied from zero (coastal dune vegetation) to more than 70% (desert scrub and desert sand). The area occupied by the seven LULC classes with CVs higher than 50% represented about half of the total classified image.

Based on the results of the error matrix and CVs, and the similarities in geographic location and floristic structure of the vegetation types, the original 26 LULC classes were reduced to 19. LULC classes were merged as follows: (1) open short forest with pine-oak forest; (2) short and medium tropical deciduous forests; (3) short and medium tropical evergreen forests; (4) introduced grassland with cultivated grassland; (5) mesquite forest, sand desert vegetation, halophytic vegetation and desert scrub. This procedure resulted in an increase in the CVs. The overall correspondence value increased 8% (from 46 to 54%). LULC classes with CVs above 50% represented 55% of the total classified image. CVs for individual LULC classes ranged from 7.7% (coastal dune vegetation) to 82% (desert scrub). Comparisons of the percentage areas mapped for each class between the 19-class map and the INEGI reference map can be seen in Table 1.

Again, following the criteria of geographic location and floristic affinities, a final merging of LULC classes was carried out in an attempt to further improve correspondence values. This time, cloud forest and tropical (medium) evergreen forest were merged with the tropical (high) evergreen forest, and Tamaulipan thorn scrub, subtropical scrub, submontane scrub, and short thorn forest were all merged, resulting in 14 LULC classes. In this merged classification, however, no significant improvement in overall CV (55.5%; $\kappa = 0.4696$) was apparent. The area occupied by the LULC classes with CVs above 50% increased to almost 60% of the total area, however.

DISCUSSION

The results of this mapping effort demonstrate the capability of AVHRR-NDVI time-series data for revealing seasonal changes in vegetation at the country scale for Mexico. Even though the overall correspondence value obtained was relatively low (i.e., 54%), the values for major individual LULC classes were significantly higher (82% for desert scrub, 67% for pine-oak forest, and 57% for tropical

(medium) evergreen forest). On the other hand, the three land use classes (irrigated cropland, seasonal cropland, and introduced grassland), which are estimations of the extent of human-transformed environments, registered CVs below 40%. Despite these contrasts, the results suggest that year-round phenological behavior, as revealed by the NDVI data, can be used to map general patterns in the spatial distribution of Mexico's main vegetation formations. The future deployment of new remote sensing satellites such as Terra, with the improved spectral and spatial resolutions of its MODIS sensor, will certainly enhance future mapping capabilities.

We propose the 19-class LULC classification over the 26-class and 14-class maps for the following reasons: (1) Its CVs improved significantly over those of the 26-class analysis, while the 14-class analysis showed only minor improvement; (2) vegetation formations such as cloud forest and tropical (high) rain forest need to be individually mapped given the critical role that these vegetation types have with respect to biodiversity (*i.e.*, they include "hotspots" of biodiversity); and 3) the country's different shrub formations (tropical thorn forest, Tamaulipan thorn scrub, submontane scrub, and subtropical scrub) have distinctive geographical distributions [4], that are important to map separately.

Three immediate recommendations can be made to improve classifications further using this general approach: (1) Produce improved AVHRR-NDVI data sets to produce data with reduced cloud contamination capable of reflecting better phenological processes occurring in different regions; (2) diversify reference data sources both to improve selection of training areas, and to estimate the classification performance without relying solely on correspondence analysis; and (3) incorporate ancillary data, such as elevation, to reduce misclassifications based on spectral patterns alone. We expect that further refinements of this approach will produce robust algorithms capable of producing rapid and inexpensive land cover calculations for Mexico and elsewhere.

REFERENCES

- [1] Di, L., and Hasting, D. A., 1995, Temporal stability of some global NDVI products derived from NOAA/AVHRR GVI. *International Journal of Remote Sensing* 16(18):3569-3583.
- [2] Sorani, V. and Alvarez, R., 1996, Hybrid maps: Updating Mexico's forest cartography using Landsat TM imagery and land use Information. *Geocarto International* 11(4):17-23.
- [3] Mora, F. and Iverson, L. R., 1997, Dynamic stratification of the landscape of Mexico: Analysis of vegetation patterns observed with multi-temporal remotely sensed images. *Geocarto International* 12(2):73-87.
- [4] Rzedowski, J., 1981, *Vegetación de México*. Editorial Limusa, S. A., Mexico.

Table 1. Percentage of areal extent of 19 vegetation classes using a supervised classification of time-series AVHRR NDVI data in comparison to a reference map produced using TM imagery.

Land use/ land cover class	Reference Map (INEGI)	Supervised Classification (ISODATA)
Desert scrub	27.7	23.3
Pine-oak forest	16.8	15.7
Tropical (med.) evergreen forest	4.1	4.7
Tropical deciduous forest	11.1	11.4
Tropical (high) evergreen forest	2.1	2.1
Irrigated agriculture	4.6	2.7
Tamaulipan thorn scrub	2.0	1.9
Seasonal agriculture	11.4	9.5
Natural grassland	6.0	7.3
Aquatic and semi-aquatic veg.	0.6	0.6
Introduced grassland	7.6	8.9
Subtropical shrub	1.6	2.1
Submontane scrub	1.4	2.1
Thorn short forest	0.5	0.5
Cloud forest	0.9	1.7
Savanna	0.4	0.7
Mangrove	0.5	0.7
Bare soil	0.5	4.0
Coastal dune vegetation	0.2	0.1

ACKNOWLEDGMENTS

This work was funded by the National Science Foundation, grant DEB-9711621. In addition to NSF, we would like to thank the Instituto Nacional de Estadística Geografía e Informática (INEGI) and the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) for providing the high resolution land use/ land cover map of Mexico. Thanks also to Jorge Soberón Mainero, Victor Sánchez-Cordero, and Robert Buddemeier for their important contributions to this project.